



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2011

Why be yourself? Kantian respect and Frankfurtian identification

Henning, T

Abstract: Harry Frankfurt has claimed that some of our desires are ‘internal’, i.e., our own in a special sense. I defend the idea that a desire’s being internal matters in a normative, reasons-involving sense, and offer an explanation for this fact. The explanation is Kantian in spirit. We have reason to respect the desires of persons in so far as respecting them is a way to respect the persons who have them (in some cases, ourselves). But if desires matter normatively in so far as they belong to persons, then it matters whether they really do belong to the persons who have them. Thus Kantian considerations explain why identification (or internality) is a normatively relevant category. This account is superior to others, and does not lead to reasons bootstrapping or a self-centred conception of deliberation.

DOI: <https://doi.org/10.1111/j.1467-9213.2011.707.x>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-50439>

Journal Article

Published Version

Originally published at:

Henning, T (2011). Why be yourself? Kantian respect and Frankfurtian identification. *Philosophical Quarterly*, 61(245):725-745.

DOI: <https://doi.org/10.1111/j.1467-9213.2011.707.x>

WHY BE YOURSELF? KANTIAN RESPECT AND FRANKFURTIAN IDENTIFICATION

BY TIM HENNING

Harry Frankfurt has claimed that some of our desires are 'internal', i.e., our own in a special sense. I defend the idea that a desire's being internal matters in a normative, reasons-involving sense, and offer an explanation for this fact. The explanation is Kantian in spirit. We have reason to respect the desires of persons in so far as respecting them is a way to respect the persons who have them (in some cases, ourselves). But if desires matter normatively in so far as they belong to persons, then it matters whether they really do belong to the persons who have them. Thus Kantian considerations explain why identification (or internality) is a normatively relevant category. This account is superior to others, and does not lead to reasons bootstrapping or a self-centred conception of deliberation.

I. INTRODUCTION

Following some seminal papers by Harry Frankfurt, many authors have accepted his idea that some of our desires are (to use a typical phrase) 'in an important sense our own'.¹ The recent literature abounds in attempts to explain what it means for desires to be our own in this special sense. I shall not join in these endeavours here. Rather, I shall ask what is important about a desire's being our own in this sense. Given that one has a desire, why would it matter whether one is 'identified with' it, or whether it is 'internal' to one, to use Frankfurt's terms?

Of course, in asking why being identified with desires 'matters' or 'is important', we may be asking different things. The sense in which Frankfurt holds that internal desires matter seems unproblematic. He is primarily interested in the fact that an agent *A*'s internal desires represent what is important *to A*. So there is no question as to why *A* cares about acting on these desires. However, as I shall try to show, internal desires also matter in a normative sense. I shall argue that the fact that a desire is our own in Frankfurt's sense gives rise to normative reasons, reasons we would not have had

¹ See, e.g., J. Christman, 'Autonomy and Personal History', *Canadian Journal of Philosophy*, 21 (1991), pp. 1–24, at p. 1.

if we had merely found ourselves with a corresponding ‘alien’ desire. This normative relevance of internality may not be what Frankfurtians are mostly interested in. But of course that does not show that it is not worth examining anyway.

However, arguing that ‘internal’ desires have a special normative or reasons-involving impact is not my primary aim. My main goal is to offer an explanation of this fact. Surprisingly, the explanation is Kantian in spirit, if not in letter. On some Kantian views, some desires have a particular normative import in so far as respecting them is required in order to respect the persons who have them. Given this, it is clear why ‘internality’ or ‘identification’ should be normatively relevant categories. If desires matter normatively in so far as they belong to persons, it is relevant whether a desire really does belong to the person who has it. Thus the category of ‘internality’ or ‘identification with’ desires plays a normative role. As it turns out, a Kantian view of respect for persons is a fruitful framework for explaining the normativity of Frankfurtian internality.

The primary virtue I claim for this Kantian approach is that it explains the normativity of internality without yielding certain implausible consequences which may seem inevitable. First, the account does not lead to an implausible kind of bootstrapping.² As I shall argue, people should indeed attach normative weight to facts about what their own desires are; at the same time, I think it would be absurd to claim that people can give themselves a reason to ϕ merely by forming an internal desire to ϕ . My account vindicates the former diagnosis and avoids the latter consequence. Secondly, it does not force upon us a ‘self-centred’ conception of deliberation. It respects the fact that in making decisions we usually do not worry about ourselves and the status of our desires, but about the reasons for taking various courses of action.

One remark before I start. I think that the problem I discuss arises independently of more general questions regarding the relation between desires and normative reasons. The basic phenomenon I am concerned with is the fact that having an ‘internal’ desire gives rise to certain normative reasons in a way in which merely finding oneself with an ‘alien’ desire does not. This fact should seem puzzling independently of whether one thinks that desires give us reasons in general. If one adopts the Humean view that desires ground normative reasons, one must explain what is special about those desires which are internal. If, in contrast, one denies that desires by themselves give us reasons, one has to explain, or at least explain away, the apparent fact that internal desires do seem to do just this. The usual anti-

² See, e.g., M. Bratman, ‘Intentions and Means-End Reasoning’, *Philosophical Review*, 90 (1981), pp. 252–65; J. Broome, ‘Normative Requirements’, *Ratio*, 12 (1999), pp. 398–419.

Humean strategy explains away the apparent normative weight of a desire in terms of the value of *what is desired*. But I doubt that our internal desires are necessarily directed at especially valuable ends. Nor is it clear that they have a special role in making us happy. (Sartre's famous protagonist may have the internal desire to join the Resistance, but that does not mean that doing so makes him *happy*. Of course we could define happiness in such a way that identification is an important constituent of happiness. But that would just shift the problem, for we would then have to answer the question why we should strive to be happy *in this sense*.) So anti-Humeans face a problem as well. My own sympathies are with the anti-Humean, and I shall not take great pains to frame the discussion (and present my solution) in a way that hides these sympathies. But I stress at the outset that the problem I try to solve could not be side-stepped by embracing Humeanism.

II. THE PROBLEM OF THE NORMATIVITY OF INTERNALITY

I begin with some phenomenology. When we have conflicting desires, we rarely regard the contenders as being fully on a par. By this I do not mean that they may have different strengths. Rather, I am thinking of the fact that often in cases of conflict, we as agents are clearly on *one* side of the conflict, and that only one of the desires is expressive of *our* standpoint. Suppose you want to eat less and lose some weight. You also want to go to the freezer to get some ice cream right now. Although both desires are yours in the sense of occurring in your mental life, what *you* want may be to lose weight, while the desire for ice cream may merely be 'something inside you' that moves you to act. If it turns out that you cannot resist this latter desire, you may later claim that you ate ice cream although you did not really want to do it.

An important goal of Frankfurt's work over the years has been to explain this distinction. In his terminology, the desires that represent what *we* want are 'internal', and we 'are identified with them'. In what follows, I shall adopt Frankfurt's terms, and phrases like 'where the agent stands', 'what the agent really wants', 'desires that are expressive of our identity', 'our own desires', etc., are intended to capture the same phenomenon.

There have been numerous attempts to explain what it is for a motive to be internal. Frankfurt's original account invoked higher-order attitudes; in recent years, he has added notions of caring, volitional necessity and being satisfied.³ Others have proposed additional or alternative conditions. According to Bratman, persons have long-term plans that structure their

³ See H.W. Frankfurt, *The Importance of What We Care About* (Cambridge UP, 1988), and 'The Faintest Passion', in his *Necessity, Volition, and Love* (Cambridge UP, 1999), pp. 95–107.

actions and decisions over time, and it is by being properly integrated in these diachronic structures that attitudes count as being expressive of our practical identity. Others, like Ekstrom, introduce conditions of synchronic coherence. Finally, Shoemaker and Jaworska point out that certain emotional aspects of first-order motives play an important role in defining which of them can count as internal.⁴ There are other approaches besides these.

I shall not join these attempts to analyse the nature of internality here.⁵ What I am interested in is the question of why internality *matters* normatively speaking. In discussing this question, I shall remain non-committal with regard to the question of the nature of internality as far as I can.

In this section I shall try to convince you that intuitively, internality does matter normatively. As I said, the category of internality is meant to capture where the agent *A* stands among the many motives that occupy *A*'s mind at a particular time. To say that internality matters normatively is to say that it matters where *A* stands. In certain decisions, we think that *A* should have the last word, not just any desire which *A* happens to have.

For example, we sometimes criticize people for failing to act on their internal desires. Suppose you know that a close friend of yours really wants to spend less time on his career and focus more on other things. But when you meet him today, he seems desperate and admits that he has just consented to doing a time-consuming extra project which his boss asked him to do. He says 'I just couldn't help it, I was driven by a desire to please my boss. Standing before him, it was as if I wasn't myself!'. Hearing this, you may not want just to pity his alienated state. You may want to criticize him in a normative register: 'Why did you do that? If you didn't really want to do it, you *shouldn't* have!'. Clearly you do not criticize your friend for acting against his desires. For his decision reveals that he did have a strong desire to accept the extra work. What you criticize is that he acted against his *true* desires. This suggests that facts about internality have some bearing on what an agent normatively ought to do.

Whatever reasons may be in play here, they could only be *pro tanto* reasons. You may find that all things considered, there was most reason for your friend to accept the extra work even if this went against his own desires. For instance, you may think that if he had refused to do the extra work, he would have had excessively bad feelings about it afterwards. Or

⁴ See Bratman, 'Reflection, Planning, and Temporally Extended Agency', *Philosophical Review*, 109 (2000), pp. 35–61; L.W. Ekstrom, 'A Coherence Theory of Autonomy', *Philosophy and Phenomenological Research*, 53 (1993), pp. 599–616; D.W. Shoemaker, 'Caring, Identification, and Agency', *Ethics*, 114 (2003), pp. 88–118; A.W. Jaworska, 'Caring and Internality', *Philosophy and Phenomenological Research*, 74 (2007), pp. 529–68.

⁵ I have done so elsewhere: see T. Henning, *Person sein und Geschichten erzählen* (Berlin: de Gruyter, 2009).

you may think that his other interests are really worthless, so that all in all, working overtime is the more worthwhile thing for him to do. Still, in so far as *he* did not want to accept the extra workload, there is at least some respect in which his decision deserves criticism.

Clearly, internality is not always what matters most. In fact, sometimes internality does not seem to matter at all. In some cases this is not difficult to explain. Suppose you, like me, hold that some moral requirements are categorical and have an overriding normative force. Then it follows that there are cases in which all of our desires are irrelevant, irrespective of whether they are internal. So if, for instance, you can save a child from drowning in a bathtub, there is just no relevant question as to whether the desire to do it is yours. But there is another class of cases in which the internality of our desires does not seem to matter. As I shall show, these pose more of a challenge to the theory I shall develop here. In these cases, our desires *do* matter, but their internality seems to be irrelevant: for instance, if you face a trivial choice between pasta and pizza, and you have a desire for the latter and no countervailing motive, it does not seem to be relevant whether or not this desire is 'yours'. I shall discuss this kind of case in §VIII.

The cases in which it seems most obvious that internality does matter are decisions about which life to choose, which job to take, whom to marry, etc. Such decisions concern what might be called *personal ends*. In decisions about personal ends, we are not willing to act on *any* desire we happen to have. Just because there is a strong impulse with which you find yourself, you do not consider yourself to have a reason to, say, move to a new continent or get a divorce. If anything, such impulses irritate you and maybe initiate self-reflection on your part. ('Do I really want *that*? Why? Since when?') But if you judge that *you* really want to move to a new continent or get a divorce, this may well affect the reasons you have. (As I shall show below, it is not clear that this internal desire *is* a reason to move to a new continent or to get a divorce. Rather, I shall claim that this desire gives rise to certain higher-order reasons that affect which considerations you should take into account and how you should weigh this desire against other impulses you have.)

My aim in this paper is to explain what it is about internal desires that gives them this normative impact. One condition of adequacy for such a theory is that it has to explain the scope of the relevance of internality. Why does internality matter in just those cases in which it does, and why are there (seemingly) cases in which it does not matter at all?

One might object that I approach these matters in much too abstract and general a way. As I have said, there are many substantive accounts of what constitutes internality. So it is natural to suppose that each of them may

yield its own specific account of why internality matters normatively. If, as Bratman holds, internality is a matter of being integrated in stable diachronic intentional structures, then maybe internality matters because diachronic stability matters. Or if internality depends on second-order desires, maybe we have reason to pursue internal desires because the second-order desires provide us with additional reasons (as Humeans may think).

Faced with these proposals, I want to point to the following *desideratum*. I think we should try to develop an account of why internality matters *as such*. My intuition is that internality matters *because* it matters where we stand. If stability and coherence determine where the agent stands, that is part of what makes them important. So we should not defer to these substantive accounts to explain the normative relevance of internal desires. At least, my hypothesis in what follows is that internality matters as such, and that this requires explanation.

III. INTERNALITY AS REASONS-RESPONSIVENESS

That said, there is one particular account of internality which I shall discuss, because it may seem to have a special bearing on my question. Some authors, most notably Watson and Raz, claim that desires are ‘our own’ if and only if they are *responsive to (beliefs about) reasons or values*.⁶ To give a famous example, Quinn’s ‘radio man’ feels alienated from his persistent desire to turn on radios *because* he fails to see anything good in doing it.⁷ One may think that this kind of account yields an explanation of the normative relevance of internality. On this account, if you act against an internal desire, you *ipso facto* act against what you believe is valuable, or contrary to the reasons you think there are. Surely there is something wrong with this. (In fact it is not clear that there really is something wrong with this. Recent debates about bootstrapping have revealed that it is not clear that beliefs about reasons or value give us reasons themselves. Broome holds that these beliefs merely ground *normative requirements*; others deny even this.⁸ For the sake of argument, I shall grant that somehow these beliefs are indeed normatively relevant.)

⁶ G. Watson, ‘Free Agency’, *Journal of Philosophy*, 72 (1975), pp. 205–20; J.W. Raz, ‘When We Are Ourselves’, in Raz (ed.), *Engaging Reason: on the Theory of Value and Action* (Oxford UP, 1999), pp. 5–21.

⁷ W. Quinn, ‘Putting Rationality in its Place’, in his *Morality and Action* (Cambridge UP, 1993), pp. 228–55.

⁸ On bootstrapping, see Broome, ‘Normative Requirements’; N.W. Kolodny, ‘Why Be Rational?’, *Mind*, 114 (2005), pp. 509–63.

Does this account answer the question? No. The problem is that responsiveness to reasons or value is not sufficient for internality. There is a sense in which even many external impulses are responsive to the value we find in things. If you find yourself with an ‘alien’ desire for ice-cream, this is usually not at all like the odd desire to turn on radios in Quinn’s example. It is not an unintelligible spasm of motivation; it is responsive to the appealing things you see in ice-cream, and very much so. Still, it is a response that you do not recognize as yours. Something similar applies in the case of your workaholic friend. It is not as if his attachment to his job is wholly unintelligible. Rather, it is clear that he finds value in pleasing his boss, in having a successful career, etc. Again it is just that these are not the kinds of considerations to which he means to be responsive. These cases show that responses to reasons need not be internal. In general, desires can be responsive to reasons and/or values even though we experience them as alien.

One may want to object to these claims. Of course, it would be question-begging to say that the desires I described cannot be responsive to value *because* they are not internal, i.e., because we experience them as mere goings-on inside us. On the face of it, my examples show that there is a sense in which our conative states can be responsive to value *and* external, something we find ourselves with. You may insist that responsiveness to reasons requires that the subject must be in charge. But then you *define* the right kind of responsiveness or value in a way that presupposes an understanding of what it is for the agent to be in charge, i.e., of internality.

But there is another objection. In the examples I gave, the subjects do not consider themselves to have *most* reason to eat ice-cream or to accept an extra job. Nor do they find *most* value in these things. What they think most valuable is losing weight or having a private life beside their job. So maybe we should claim that a desire is internal if and only if it is responsive to the subject’s believing that there is most value in its object, or most reason to pursue it?

This would be an implausibly rationalistic understanding of internality. Surely it is not the case that whenever we are identified with a desire to ϕ , we take ourselves to have most reason to ϕ . We may take our desire to be a mere inclination and still be fully behind it. As Watson himself stresses in later work, we may even have an internal desire to ϕ even though we judge that we have most reason *not* to ϕ .⁹ For instance, it seems that even if one admits that one really should not smoke, one still can be fully behind one’s desire to do so. This desire may be irrational, and in acting on it one is akratic (in the sense of not acting in accord with one’s judgement about

⁹ See Watson, ‘Free Action and Free Will’, *Mind*, 96 (1987), pp. 145–72, at p. 150.

what is best). But one's desire is surely no less wholehearted for that. Since the desire is (arguably) not immoral or excessively imprudent, there even seems to be some reason to respect it. If Alice tells us 'I know I shouldn't smoke, but I want to, so back off', this gives us a reason to respect her will. Intuitively, things would be different if Alice merely confesses to have a desire which she considers alien. (I shall come back to this kind of case below.)

To sum up: responsiveness to (beliefs about) reasons or value is not sufficient for internality, and responsiveness to (beliefs about) what is best or what we have most reason to do is not necessary for internality and the normative weight of internal desires.

IV. INTERNALITY AND THE VALUE OF AUTONOMY

Before I present my own account, it may be helpful to consider another alternative approach. To many, this approach may seem to be the obvious solution. As is well known, internality is most often discussed in recent debates about *personal autonomy*. Starting with Frankfurt, the general thrust of these recent discussions has been to move away from the 'impersonal' Kantian notion of autonomy and towards a view on which autonomy is a matter of pursuing concerns which are one's own. On this view, autonomy requires that one should pursue one's internal motives.

If pursuing one's internal desires is required for personal autonomy, then maybe this is the clue to the importance of internality? Surely autonomy is a valuable thing, and we have reason to strive to be autonomous. But then it seems that we can say that we have reason to act on our internal motives *because* this is required for personal autonomy. So in the example above, you would criticize your workaholic friend for not being autonomous.

There is something to this proposal. As I shall argue below, there is indeed a sense in which the importance of internality is grounded in our status as beings who are capable of making free and responsible decisions and who deserve respect. But in the present form, the proposal is still unsatisfactory. 'Autonomy' can mean different things (even if we focus on uses of 'autonomy' that apply to persons). Here are just two important senses: on the one hand, autonomy is the general status I have just invoked, i.e., the status of a being who is *capable* of self-government. I fully agree that autonomy in this sense is valuable in itself. However, acting on one's internal desires is obviously not required for being *capable* of self-government, i.e., for autonomy in this first sense. It is required for being autonomous in a second sense, the sense in which being autonomous means

manifesting the general capacity.¹⁰ In this sense, we can say that one is autonomous *in acting* in certain ways.

Since acting on one's internal desires is only required for *manifesting* autonomy, the proposal would have to be that this latter fact explains why we have reasons to act on internal desires. I do not think that this explanation is persuasive. While it is true that the general capacity of autonomy is valuable and worth cherishing, it is considerably less clear that the fact that a particular action would manifest that capacity gives us a reason to perform it. When was the last time you decided to do something *because* doing it was autonomous? To my ear, it sounds fetishistic to say that one did something for the reason that it was a way to manifest one's autonomy. (I am not here assuming the Kantian framework in which doing something because it is autonomous is roughly equivalent to doing it because it is in accord with the moral law. In such a framework, it may be that we can make sense of the idea of doing something because it is an autonomous action.) But if the fact that an action manifests one's autonomy does not give us a reason to do it, the fact that acting on internal desires is necessary for (or constitutive of) such a manifestation cannot explain why internality matters normatively.

To be sure, we do criticize people for not being autonomous with regard to certain actions of theirs. But in doing so we do not mean that the reason why they should have performed the action was that it would have been autonomous to do so. In this regard, autonomy is analogous to rationality. We do criticize people for failing to do what they rationally ought to do; but that does not mean that the reason they had for doing it was that the action was rational (cf. Kolodny, p. 545). Rather, an action is rational in virtue of responding to other, more substantive reasons (or beliefs about such reasons). When we call an action rational, we do not usually mean to provide a further, or more fundamental, reason there was for doing the action. When we criticize irrationality, we criticize failure to respond to certain other reasons, reasons which are not facts about rationality. I think that when we criticize our workaholic friend for not being autonomous, our criticism has similar features. The reason that we think he has failed to respond to is not that the action he failed to perform would have been autonomous, but rather that this action would have been just *what he really wanted to do*. A failure of autonomy precisely is a failure to act on this latter kind of consideration. Thus even if acting on one's internal desires is acting autonomously, that could not explain why facts about internal desires give rise to reasons.

¹⁰ Cf. J.W. Feinberg, *The Moral Limits of the Criminal Law*, Vol. 3: *Harm To Self* (Oxford UP, 1986), pp. 27ff.

V. DESIRES, REASONS AND RESPECT

A couple of recent authors claim to find in Kant an explanation of why our desires matter normatively. I shall try to show that a plausible version of this account of the import of desires yields an account of the import of *internal* desires. In fact, this account will give a plausible explanation of why the concept of internality is needed in the first place. Thus a Kantian framework explains the (normative) importance of our internal desires.

Authors like Korsgaard and Wood claim that Kant's argument for the Formula of Humanity (in section II of the *Groundwork*, IV 426) relies on a claim about the connection between the value of what we desire and our own value as subjects of the desire.¹¹ This interpretation of Kant is controversial and widely discussed, and I shall not enter this discussion here. But whether or not Korsgaard and Wood offer a correct reconstruction of Kant's argument, I do think that they correctly identify a central Kantian insight. Simply put, the important thought is this: sometimes, desires give rise to normative reasons because neglecting or disregarding them is a way of disrespecting the persons who have them.

The best development of this general line of thought that I know of is due to another (broadly) Kantian author, Stephen Darwall. He gives the following example:

Parents may reasonably treat the desires of their very young children as having no intrinsic weight. The mere fact that their one-year old daughter doesn't want to eat a healthful food is not a reason *in itself* for them to desist in efforts to get her to do so.... When, however, she is full grown, her desires do acquire intrinsic weight for them.... Suppose you are a middle-aged daughter whose parents are trying to get her to eat 'her broccoli'. You might well think that the fact that you don't *want* to eat your broccoli is a reason why you should be allowed not to do so (thank you, very much!), and, indeed, that it is a reason why you should *allow yourself* not to do so. Simply deferring to your parents' wishes at this point might well be a failure to respect yourself as an equal.¹²

Darwall observes that the normative relevance of desires seems to depend on the status of the subjects of the desires. If a grown-up person does not want to eat broccoli, we have reason to accept this. If the subject of the

¹¹ See C.W. Korsgaard, *Creating the Kingdom of Ends* (Cambridge UP, 1996), and *The Sources of Normativity* (Cambridge UP, 1996); A. Wood, *Kant's Ethical Thought* (Cambridge UP, 1999).

¹² S.L. Darwall, 'Desires, Reasons, and Causes', *Philosophy and Phenomenological Research*, 67 (2003), pp. 436–43, at p. 442. For a more extensive discussion, see Darwall, 'Because I Want It', *Social Philosophy and Policy*, 18 (2001), pp. 129–53.

desire is a young child, things may be different. Furthermore, Darwall explicitly mentions the category of *respect* in this connection. If parents try to make their adult daughter eat broccoli against her will, this shows a lack of respect. Finally, he claims that people owe this respect *to themselves* as well. If a grown-up daughter ignores her own wishes as soon as her parents (or her neighbours, or anybody) tell her to do so, this is a lack of self-respect.

In recent years, Darwall has done much to explore the sense in which respect gives both the parents and the daughter a reason to ascribe normative weight to her desire. On his view, the reason is of a distinctive ‘second-person’ kind, namely, a reason which is ‘rooted, not in the objects of any desire (the value of a possible state of the world), but in claims [the daughter] has the authority to make as a free and rational will’.¹³ The analysis of second-person reasons is complicated and I cannot give a reconstruction here.¹⁴ But I can give an analogy to illustrate the idea: on Darwall’s view, we have reasons for respecting a person’s choices in something like the way in which, e.g., a soldier has a reason to obey a commanding officer’s order. This reason has nothing to do with the valuable state that will result from obeying. In fact, the soldier is not really obedient if he does as he is told only because he thinks the commanded action worthwhile for independent reasons. Rather, he owes compliance to the officer. The reason is rooted in the officer’s authority to demand this compliance.

I think the analogy brings out an important aspect of respect. Suppose the parents decide to stop importuning their daughter, but they do so only for the reason that they have changed their mind about the value of eating broccoli. If they allow their daughter not to eat broccoli only because they judge that not eating broccoli is the better action after all, or because this will avoid a long argument which will ruin the day, they still fail to respect her. Alluding to Bernard Williams, Darwall (‘The Value of Autonomy, p. 268) says that the parents’ practical reasoning ‘would already involve one or two thoughts too many’. Respecting their daughter requires the parents to back off *because she wants it* – more specifically, because of ‘the authority that persons have to demand, within certain limits, that they be allowed to make their own choices for themselves’.

Thus what respect requires is precisely that the parents should back off although they fail to see anything good in not eating broccoli. We treat certain claims as reason-giving in a way which cannot be reduced to agent-neutral considerations about valuable states of the world. As a matter of fact,

¹³ Darwall, ‘The Value of Autonomy and the Autonomy of the Will’, *Ethics*, 116 (2006), pp. 263–84, at p. 280.

¹⁴ See Darwall, *The Second-Person Standpoint: Morality, Respect, and Accountability* (Harvard UP, 2006).

respect does not really require us to treat people's desire to ϕ as a reason for them to ϕ . In the case of the wholehearted smoker, if Alice says 'I know I shouldn't smoke, but I *want* to, so back off', her aim could not be to refer to an overriding reason which counts *in favour of her smoking*. In that case, what she says would contradict the first part of her statement. Rather, it seems that what she has given is a reason which operates on a different level, a reason which does not outweigh the ordinary reasons against smoking but, as it were, puts them aside.

Following Raz, I shall call these reasons *exclusionary reasons*.¹⁵ Exclusionary reasons are second-order reasons. So if Alice's ϕ ing is supported by an exclusionary reason, this reason does not pertain directly to the good aspects of her ϕ ing. Rather, it is a reason against taking such first-order reasons into account, criticizing actions and disputing decisions in the light of them, etc. It is easy to see why respect would be a source of such exclusionary reasons: if, for instance, we insistently object to Alice's decision to ϕ because we think that she ought not to ϕ , we treat her decision as if it was *ours* to make. Of course, there are many subtle distinctions that are crucial here. Offering advice to the effect that Alice should not ϕ is usually legitimate. (However, if ϕ ing is not immoral and the reasons against it are fairly obvious, it can indeed be disrespectful to give a stranger a lecture about the negative aspects of ϕ ing.) But to keep on insisting even after Alice has heard our advice often amounts to a lack of respect. (Of course, this is not to deny that considerations of respect can be overridden by other reasons.)

Anyway, to respect Alice's authority to make her own choices is to accept these choices unless there are very strong reasons against it. In his description of the broccoli case, Darwall invokes second-order reasons as well. He says that the daughter's desire should count as a reason 'why [she] should be allowed not to [eat broccoli]'. Clearly, a reason for *allowing* people to ϕ need not be a reason why they ought to ϕ . Rather, it may be a reason not to *object* to their ϕ ing although one thinks that the balance of first-order reasons counts against it.

Thus I think part of Darwall's proposal can be captured by saying that respect for persons gives us an exclusionary reason to respect their decisions, a reason to abstain from objecting to these decisions although we would reject them in the light of the relevant first-order reasons. The fact that these reasons are, to quote Darwall (*The Second-Person Standpoint*, p. 18), 'not appropriately weighed against otherwise competing considerations' seems to

¹⁵ One of Raz's own examples is precisely the case of obeying an order, i.e., the case that I have used to illustrate the irreducible character of Darwall's 'second-person reasons': see Raz, *Practical Reason and Norms* (London: Hutchinson, 1975), p. 38. Darwall mentions exclusionary reasons in *The Second-Person Standpoint*, p. 18.

me to be a way to capture Kant's idea that persons have a claim to being treated as ends in themselves.

However, as Darwall rightly stresses here, being an exclusionary reason is not sufficient for being the kind of second-person reason which is connected with respect. Mere exclusionary reasons, being agent-neutral, do not suffice to capture the kind of authority that the agent has. In the next section, I explain that respect for persons is a source of exclusionary reasons which are not agent-neutral. In the first instance, it gives other persons reasons not to try to make decisions for the agents. From the perspective of the agents themselves, respect gives rise to reasons that *derive* from those agent-relative reasons of others. Respect does not exclude any first-order reasons from the agent's own deliberations. But it requires them not to let others treat their own decisions as something that is up to those others and answerable to their view about what to do.

Much as I agree with Darwall's account of respect for persons and their wishes, there are different ways to spell out the details. As I shall argue now, Darwall is not clear on one point. Reflection on this point about respect and the reasons it gives us displays the germ of an explanation for the special weight of internal desires.

VI. RESPECT AND INTERNALITY

As I have explained, Darwall thinks it is important that the parents should back off just because their daughter does not want to eat broccoli. I do think that this claim is correct, and I have said something about how it should be interpreted. However, there is still need for some careful interpretation of one particular point. According to Darwall, respect requires us to ascribe 'intrinsic' normative weight to certain desires. To this, I say 'Yes and no'. I agree that if the parents are to respect their daughter, they must not merely do so for the reason that letting her have her way is a means to achieve some other end, say, having a peaceful afternoon. In this regard, we can indeed call the desire's relevance intrinsic. Still, there is a sense in which respecting the desire does serve some other end, albeit an end to which it is not related as a means. The parents must ascribe weight to the daughter's desire *in order to respect her*, where respecting the desire is not a means towards respecting her but *constitutive of* doing so. Respecting her wishes is, in the context under consideration, a way of respecting the daughter.

In this sense, it would be misleading to regard the desire as important in itself. To elaborate: respecting the daughter's desire on account of *its* alleged importance would be to miss the proper object of respect. Suppose the

parents think that the desires of adult persons are among the entities in the world which are intrinsically normatively relevant and reason-giving. Whenever they find a desire in an adult, they take this conative state to have normative significance. Clearly, they can do so without respecting their daughter as a person, and without granting her the right to make her own decisions. Suppose the daughter says ‘Listen, I do not *want* to eat this’, and the parents turn to each other and say ‘Uh oh, here we have an adult person’s desire – that changes everything’. Then, turning to their daughter, ‘It’s OK, mom and dad have just decided that you are allowed not to eat the broccoli’. Obviously, the daughter still has reason to complain that *she* plays no role in the parents’ practical reasoning, and so she may still not feel that she has been properly respected. She wants her parents to back off not because they somehow cherish adult people’s desires, but because it is what *she* demands. That is, her desire has normative significance in this context in so far as, in this context, respecting it is required in order to respect *her*. The reason why the parents should not force their daughter is that this would be to disrespect her, given what she desires.

Thus I think that in the case above, the reason which the parents have is that their daughter as a person deserves a certain kind of respect. Of course, the fact that she does not want to eat broccoli does make a difference, for it affects what, in the case at hand, respect for persons requires. We could say, using a term from Wolterstorff, that a person’s desires are among the factors that influence which actions have ‘respect–disrespect import’.¹⁶ In this way, they can affect what we ought to do without being the source of this ‘ought’.

Thus when respect for persons requires us to attach normative weight to their desires, then the *object* of our respect is still the person. We are trying to respect *them* by treating their desires in certain ways, even if we fail to see anything good in these.

But if this is right, the relation between the person and the desire turns out to be important. For if desires are normatively relevant only in so far as they speak for persons, then of course it is important whether they *do* indeed speak for the persons who have them. So once Darwall’s approach is clarified in this way, it can be seen how Frankfurtian considerations of internality can be relevant for Kantian considerations of respect for persons. Since we do not regard all of *A*’s desires as having the authority to speak for *A*, respect for *A* requires us to be discriminating. If only some of *A*’s desires speak for *A*, these desires have a special claim to be respected as *A*’s.

To take a closer look, suppose you want to work on a paper, but you find within yourself a contrary wish to go out for a drink. The former desire is yours, whereas the latter is a nagging urge which you do not recognize as

¹⁶ See N. Wolterstorff, *Justice* (Princeton UP, 2008), p. 296.

speaking for you. Alice, a good friend who knows you very well, realizes that you are about to be overcome by the desire to go out. Being your friend, she has good reason to believe that this is not really what you want. What it does it take for her to respect you and your will? Does she do it by just accepting or encouraging *whatever* desire turns out to be effective in moving you to action? It seems not. Unlike the parents in Darwall's broccoli example, it seems that Alice would not be disrespectful in trying to make you stay, not even in doing so rather insistently. This would not be disrespectful, in so far as she would not do it because she thinks it is better for you to stay, but rather because she wants to promote your true desires. (As a matter of fact, she might even think that it would be better for you to go out for a drink and leave your work for a while.)

Thus respecting persons as persons requires being sensitive to the difference between their internal desires and mere external urges or impulses. Since internal attitudes constitute a person's will, they have a special role in determining what it takes to respect that person's will.

VII. RESPECTING THE INTERNALITY OF ONE'S OWN DESIRES

I am now in a position to describe how respect for internal desires operates in the first-person case. As Darwall says, the daughter seems to fail to respect herself as a person if she simply does as she is told even though she does not want to. So we seem to have reason to respect our own internal desires. But how exactly is sense to be made of this idea? As I shall show now, this can be done by using the ingredients which I have assembled in the two previous sections.

I have previously pointed to the dangers involved here. The claim that we must even respect our own internal desires in order to respect ourselves as persons may seem to lead to absurd consequences. First, it may seem to lead to an objectionable kind of *bootstrapping*, *viz* to yield the consequence that we can give ourselves a normative reason to ϕ by forming an internal desire to ϕ . This would be absurd. If we do not have a reason to ϕ but still form an internal desire to ϕ , this does not make it any more reasonable to ϕ . If anything, we have a reason to get rid of this irrational desire.¹⁷ Secondly, the account may seem to imply that deliberation should constantly be concerned with features of one's own motivational states. This would be implausible too; as Watson ('Free Agency', p. 219) stresses, our deliberation is usually not that self-centred.

¹⁷ Cf. Bratman, 'Intentions and Means-End Reasoning'; Broome, 'Normative Requirements'.

Fortunately, my account captures the sense in which we ought to respect our own internal desires without having any of these absurd consequences. I shall begin by explaining in general outline how I propose to think about the requirements of respect for one's own internal desires. As I have indicated, the idea is that respect for internal desires gives rise to *agent-relative higher-order reasons*. First, these reasons are higher-order, in the following sense: having an internal desire to ϕ is not a reason which counts directly in favour of ϕ ing; rather, it is a reason (not) to take into account certain first-order reasons for or against ϕ ing. Secondly, these reasons are agent-relative in the sense that they are primarily reasons *for others* not to take first-order reasons into account in a certain way. Specifically, these are reasons for others not to relate to A 's decision to ϕ merely on the basis of their own view of whether A has most reason to ϕ .

These two considerations suggest an account of what respect for persons' internal desires requires of the persons *themselves*. First, respect for my internal desire to ϕ does not give me a first-order reason to ϕ . Secondly, nothing about the account requires us to hold that respect for this desire gives *me* a higher-order reason to disregard my judgement about the balance of first-order reasons. Surely, if I change my mind about ϕ ing on the basis of my own judgement about reasons for ϕ ing, acting in accordance with this new decision does not amount to lack of respect. Thus I think the right way to spell out requirements of self-respect is to explicate them in terms of what I have said about the authority of others. Just as others do not have the authority to make A 's decisions for A , A is required not to *grant* them this authority. So respect for A 's own desires does not allow A to leave first-order reasons out of consideration; instead, it requires A not to abstain from ϕ ing just because *someone else* thinks A should not ϕ .

This seems to restrict the normative relevance of internality to interpersonal contexts. However, this is not the end of it. I am using the Frankfurtian distinction between internal and external motives precisely so as to allow that even motives and thoughts that occur within A may not really belong to A . Given this, the account just sketched entails that internality also has a considerable intra-personal relevance. Just as self-respect requires me not to let others decide what I am to do, it requires me not to let 'external' motives and appetites which I do not recognize as mine determine what I am to do. Just as letting other persons decide manifests a lack of self-respect, so succumbing to these motives and appetites against one's will can be a failure to respect oneself. Moreover, some external desires are responsive to reasons and values. In these cases, it can indeed happen that the fact that a desire is internal has a normative significance which seems to override some first-order reasons. But this is not because internal desires do give rise to

first-order reasons after all. Rather, it is because they give rise to a higher-order reason to disregard our alien *responses* to the relevant first-order reasons. For instance, it may be legitimate for you to disregard the many advantages of a certain job simply because you find that what *you* want is something else, whereas the desires that respond to considerations of safety or profit are alien impulses inside you. This does not mean that facts about your internal desires are on a par with, and weighed against, the first-order reasons in favour of the job. It means that if you acted on your desire for safety and profit, this would be a failure to respect your own will, and that you thus have a higher-order reason not to do so.

This way of spelling out what respect for one's own will requires does not only give the right result in Darwall's broccoli case. It also avoids the unwelcome consequences sketched at the beginning of this section. First, it avoids bootstrapping. It does not entail that by forming an internal desire to ϕ we can bootstrap a reason to ϕ into existence. For having such a desire merely gives rise to a reason not to let certain factors influence the decision to ϕ . So for agents themselves, it is never reasonable to weigh ordinary first-order reasons against reasons to respect their desires directly. That is, respect never requires them to say 'Well, now that I think about it, I really should not ϕ – but on the other hand, I have decided to do so, and I must respect that'. For the same reasons, my account does not entail that your internal desire to ϕ gives you a second-order reason not to take into account the first-order reasons for and against ϕ ing. It is only when you revise your decision in response to the mere fact that *others* object to your ϕ ing, or under the pressure of inner forces which you do not recognize as part of yourself, that you fail to respect yourself.

Secondly, the account does not entail a self-centred conception of deliberation. In practical deliberation, you are engaged in the process of *determining* what you want in the first place. Clearly, doing this cannot yet be a matter of respecting what you want. In fact, whenever you ask whether you should ϕ , you are concerned exclusively with first-order reasons for and against ϕ ing. It is only when you ask yourself whether you should abstain from ϕ ing because some person or some conflicting desire tells you to abstain that the question of what you really want becomes relevant.

To sum up: my account does not lead to bootstrapping or self-centredness. Our internal motives do not give rise to reasons which from our own perspective can be weighed against first-order reasons for or against doing what we desire to do. Respect for our own desires merely requires us not to let others, or alien motives, determine a different course of action for us. We are free to change our minds in the light of reasons that others may offer us; thus we are free to follow advice. But we are required not to grant others the

authority to decide for us (not even in tough cases in which we might like to have our choices be made for us). Of course, it is hard to draw a sharp line between those ways of changing one's mind which are and those which are not compatible with exercising the authority to decide for oneself. The line between following advice and succumbing to power, temptation or rhetoric can hardly be drawn with precision. But the vagueness of this distinction does nothing to cast doubt on its importance.

VIII. THE LINE BETWEEN THE INTERNAL AND THE EXTERNAL

Here is another problem for my account. As I said in §I, sometimes desires matter even though they are not internal. This fact by itself is not a problem. Nothing in my account forces me to deny that desires can give rise to reasons in other ways besides the one discussed here. But there seems to be a related fact which is problematic. Some desires seem to matter *in terms of respect* even though they are not internal. In a different version of Darwall's broccoli case, suppose the daughter's desire not to eat broccoli is a mere aversive impulse. Intuitively, the parents would have acted just as disrespectfully in ignoring this impulse. That is, it seems that even reasons of respect do not seem to attach exclusively to internal desires.

Clearly, this is a problem. For if the latter diagnosis were correct, we would still lack an explanation of the special weight of internal desires. But fortunately, this problem can be solved. I think that although it may seem that external desires matter in terms of respect, in all relevant cases the desires are not really external after all. This requires some explanation. It should be noted first that the distinction between internal and external desires *as it is usually presented* does not seem to be exhaustive. For instance, my desire to have milk for breakfast does not seem to satisfy any demanding criteria for internality. We may assume that I do not have any higher-order desires with regard to it. Nor is it part of any long-term plans and policies, or associated with attitudes of 'caring'. That, however, does not mean that this desire is on a par with the 'alien' urges and drives which are usually cited as examples of external desires. It is just an ordinary desire of mine. It is neither an 'outlaw' nor an 'intruder' in my psychology, nor does it manifest a special wholeheartedness.

I think there are two ways to conceive of the status of desires like this one. First, we might say that my desire for milk is internal even though it does not satisfy any of the demanding conditions of internality that seem to be called for in other cases. Taking this option, we would have to claim that standards for internality are not invariant, but are sensitive to other factors.

A second option would be to stick to a narrow conception of internality: we would deny that my desire for milk should be counted as internal. For reasons of parity, however, we should not claim that the desire is external either. Rather, we should treat the distinction as not exhaustive.

I think the difference between these two strategies is mainly verbal. The former has certain intuitive advantages, though. First, it would allow that the distinction between internal and external desires is exhaustive. Secondly, it would also allow me to stick to my Kantian claim that reasons of respect arise exclusively in connection with internal desires. I would then have to argue that in the case above, the daughter's desire not to eat broccoli can plausibly be regarded as internal, i.e., as on a par not with alien urges but with my desire to have milk for breakfast. On the second way of drawing the line between the internal and the external, my strategy would have to be different. I would have to claim that respect for persons requires others to respect their desires as long as these are not external. Admittedly, this move would entail that I have formulated the problem of this paper inaccurately. What I explain is the special normative weight of the distinction, not between internal and external desires, but between external and non-external desires (many of which are internal). But again this seems to be a verbal difficulty.

Anyway, both strategies would require me to say something more about where the line between the external and the internal (or between the external and the non-external) is to be drawn, and both would force me to explain why I think that the daughter's aversive impulse does not count as external. Here I take my central clue from Frankfurt's later writings. To avoid the notorious regresses faced by his theory, Frankfurt allows that one's desire may be internal simply because one *does not distance oneself* from it.¹⁸ Thus internality can derive from the mere absence of contrary attitudes (on the same and higher levels). I shall not here try to judge whether this proposal solves the specific problems of Frankfurt's theory. However, it does seem to be promising in dealing with desires like my desire for milk. We could say, then, that as long as a desire is not in tension with other desires, it counts as the person's own (or, on the alternative treatment, as non-external). On this view, it is only when there is a conflict of desires (on the same or different levels) that we need to be more discriminating and look for further conditions to help us decide which desires are a person's own.

This kind of approach does not only seem to me to be plausible considered in itself. Furthermore, it dovetails very neatly with my Kantian theory of the normative relevance of internality. If desires matter normatively in so

¹⁸ Frankfurt, 'The Faintest Passion', in his *Necessity, Volition, and Love* (Cambridge UP, 1999), pp. 95–107.

far as respecting them is a way to respect the persons who have them, it is clear why cases of conflict require special measures. While our *default* attitude is that persons' desires deserve to be respected as theirs, conflicting desires yield conflicting requirements of respect. If your desires are in conflict, it becomes a critical question what it takes to respect your will. This is why in cases of conflict we start looking more closely at the relation between the person and each of the conflicting desires, trying to find a rationale for treating only one of these as the true expression of the person's own will.

On this view, I misdescribed the variation of Darwall's case at the beginning of this section. I should not have said that the daughter's desire is not internal; rather, I should have said that even though it is a mere aversive reaction that she finds herself with, there is in the case at hand no reason to suppose that it does not belong to her. The aversion may come as a surprise, but that is no reason to regard it as alien. (As a matter of fact, I think that if we counted every spontaneous conative attitude or impulse as 'external until proven innocent', we would face an implausible proliferation of cases of alienation.) This explains why the parents would be disrespectful in neglecting the daughter's aversion. However, if they recognize that she is experiencing a conflict, and that the aversion is affecting her behaviour against her own will, then what they do in neglecting the aversion may be legitimate.

Hence the Kantian account can explain why it may appear that there are cases in which external desires give rise to reasons of respect. This appearance arises because special standards of internality apply only where the question of what the person's own will is becomes critical.

IX. CONCLUSION: WHY INTERNALITY MATTERS

In this paper I have tackled the problem of saying what is important about internal desires as such, and I have presented an account that is Kantian in spirit (if not in the letter). Of course, one could argue about how much of Kant's original views this account really contains. But it is an account which rests on the idea that persons are ends in themselves, in the sense that their desires and decisions deserve to be respected in a way which cannot be reduced to the first-order reasons which pertain to what they desire or decide to do.

The account satisfies the desideratum of being largely independent of any specific substantive account of the nature of internality. It does not reduce the relevance of internality to the independent value of any of its constituents. But there is a further advantage which I consider even more

important. The account I have offered does not start from the intuitive importance of internality and then try to give a rationale or ground which is designed to vindicate it. Rather, I have started from a general account of why a person's desires matter, an account which does not mention their internality or externality. On this general account, desires can make a normative difference because persons have a claim to respect as free and responsible beings. From these general considerations I have been able to derive an explanation of the importance of the internality of desires. I argued that if *that* is why desires matter in general, then the specific relation between the person and the desire must matter as well. Instead of taking for granted the fact that we distinguish between those desires which we merely have and those which are ours, I deliver a normative account of why we should do that kind of thing. In a way, a Kantian account of the relevance of a person's desires shows that it is to be expected that we should make these Frankfurtian discriminations.¹⁹

Universität Zürich

¹⁹ For comments on earlier versions I would like to thank the participants of the workshop 'Psychological and Self-Alienation' in Dortmund in July 2010, especially Logi Gunnarsson, Simon Hailwood, Michael Quante and Owen Flanagan; participants at the Berlin School of Mind and Brain who provided helpful comments on this material at a talk in January 2010; and two anonymous referees.